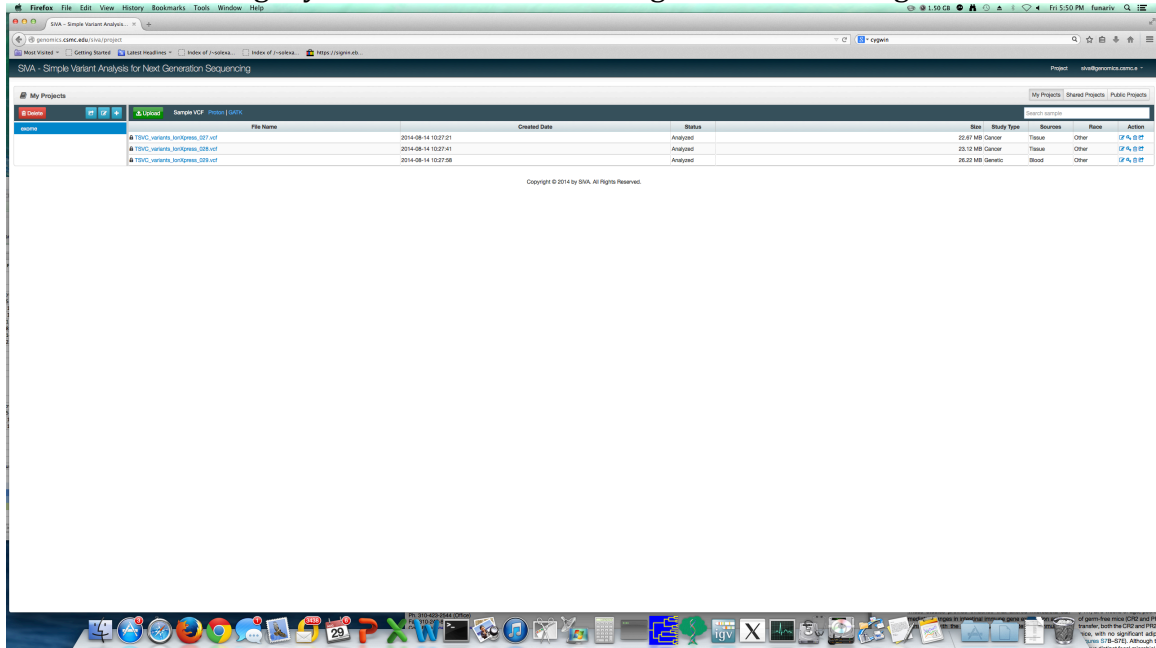


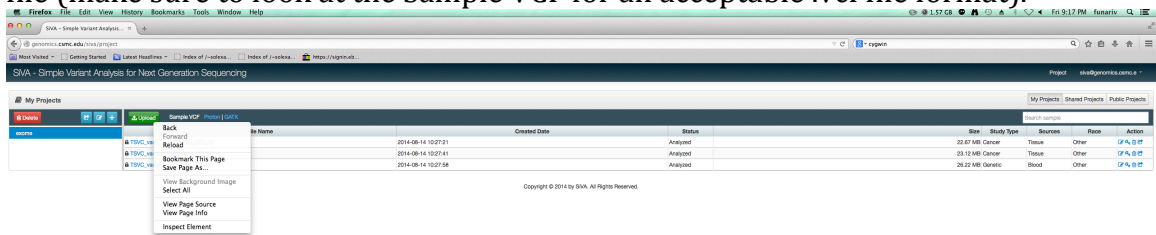
Below is an example of how one might use SiVA to interrogate exome or targeted gene panels. This is meant as an illustrated guide not a detailed manual on exome analysis, approaches or strategies. This guide assumes a basic level of genetic and genomic knowledge.

Log on to SiVA with a popular web browser (e.g.firefox) at [genomics.csmc.edu/siva/](http://genomics.csmc.edu/siva/)  
After successful login you should find something like the following screenshot:



### Upload and edit sample details:

To upload a sample, click the green **“Upload”** button to upload a FASTA file or .vcf file (make sure to look at the Sample VCF for an acceptable .vcf file format).

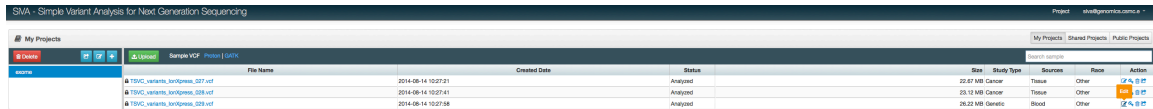


After you have uploaded your file. You will see (from left to right)

- 1) **“My projects”** directory (all projects registered to the user name),
- 2) **“File names”** (Files in the selected project directory),
- 3) **“Created date”** (Date file was uploaded)
- 4), **“Status”** (if file upload has completed will read “Analyzed”);
- 5) **“Size”** (sizes of files uploaded),
- 6) **“Study Type”** (user determined at upload either Cancer or Genetic analysis”)
- 7) **“Sources”** (user determined at upload either Tissue or Blood)
- (8) **“Race”** (user determined at upload)

9) **“Action”** (user selectable options: **“Edit”**, **“Make public”**, **“Delete”**, **“Share”** (with another user))

Select **“Edit”** to edit file name details.



SVA - Simple Variant Analysis for Next Generation Sequencing

name	File Name	Created Date	Status	Size	Study Type	Source	Race	Action
	TSDC_variants_bcrp988_227.vcf	2014-08-14 10:27:21	Analysed	22.87 MB	Cancer	Tissue	Other	<a href="#">Edit</a> <a href="#">Share</a> <a href="#">Delete</a>
	TSDC_variants_bcrp988_228.vcf	2014-08-14 10:27:41	Analysed	23.12 MB	Cancer	Tissue	Other	<a href="#">Edit</a> <a href="#">Share</a> <a href="#">Delete</a>
	TSDC_variants_bcrp988_229.vcf	2014-08-14 10:27:58	Analysed	25.22 MB	Genetic	Blood	Other	<a href="#">Edit</a> <a href="#">Share</a> <a href="#">Delete</a>

Select **“Make public”** to make file publically viewable

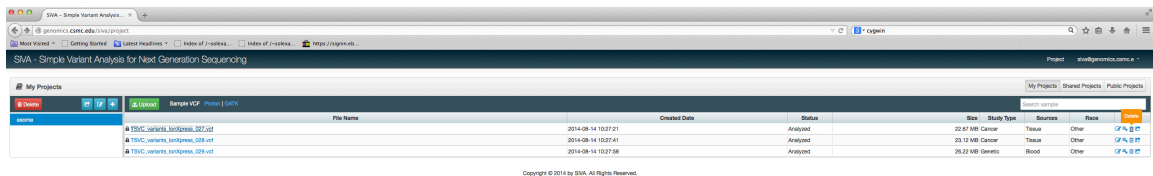


SVA - Simple Variant Analysis for Next Generation Sequencing

name	File Name	Created Date	Status	Size	Study Type	Source	Race	Action
	TSDC_variants_bcrp988_227.vcf	2014-08-14 10:27:21	Analysed	22.87 MB	Cancer	Tissue	Other	<a href="#">Edit</a> <a href="#">Share</a> <a href="#">Delete</a> <a href="#">Make public</a>
	TSDC_variants_bcrp988_228.vcf	2014-08-14 10:27:41	Analysed	23.12 MB	Cancer	Tissue	Other	<a href="#">Edit</a> <a href="#">Share</a> <a href="#">Delete</a>
	TSDC_variants_bcrp988_229.vcf	2014-08-14 10:27:58	Analysed	25.22 MB	Genetic	Blood	Other	<a href="#">Edit</a> <a href="#">Share</a> <a href="#">Delete</a>

Copyright © 2014 by SVA. All Rights Reserved.

Select **“Delete”** to delete file

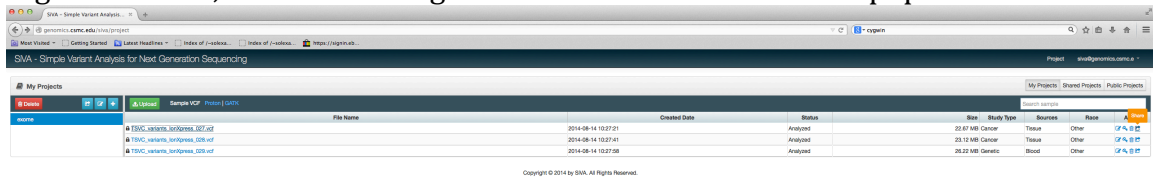


SVA - Simple Variant Analysis for Next Generation Sequencing

name	File Name	Created Date	Status	Size	Study Type	Source	Race	Action
	TSDC_variants_bcrp988_227.vcf	2014-08-14 10:27:21	Analysed	22.87 MB	Cancer	Tissue	Other	<a href="#">Edit</a> <a href="#">Share</a> <a href="#">Delete</a>
	TSDC_variants_bcrp988_228.vcf	2014-08-14 10:27:41	Analysed	23.12 MB	Cancer	Tissue	Other	<a href="#">Edit</a> <a href="#">Share</a> <a href="#">Delete</a>
	TSDC_variants_bcrp988_229.vcf	2014-08-14 10:27:58	Analysed	25.22 MB	Genetic	Blood	Other	<a href="#">Edit</a> <a href="#">Share</a> <a href="#">Delete</a>

Copyright © 2014 by SVA. All Rights Reserved.

Select **“Share”** to share with a colleague or Principle investigator who is already registered. When you share a variant results sample, please enter in the email or registered user, if the user is registered the user name will autopopulate.



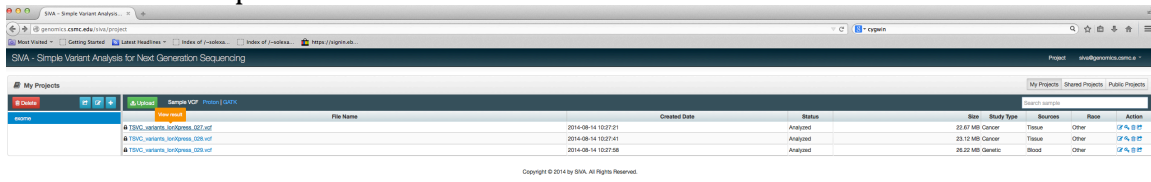
SVA - Simple Variant Analysis for Next Generation Sequencing

name	File Name	Created Date	Status	Size	Study Type	Source	Race	Action
	TSDC_variants_bcrp988_227.vcf	2014-08-14 10:27:21	Analysed	22.87 MB	Cancer	Tissue	Other	<a href="#">Edit</a> <a href="#">Share</a> <a href="#">Delete</a>
	TSDC_variants_bcrp988_228.vcf	2014-08-14 10:27:41	Analysed	23.12 MB	Cancer	Tissue	Other	<a href="#">Edit</a> <a href="#">Share</a> <a href="#">Delete</a>
	TSDC_variants_bcrp988_229.vcf	2014-08-14 10:27:58	Analysed	25.22 MB	Genetic	Blood	Other	<a href="#">Edit</a> <a href="#">Share</a> <a href="#">Delete</a>

Copyright © 2014 by SVA. All Rights Reserved.

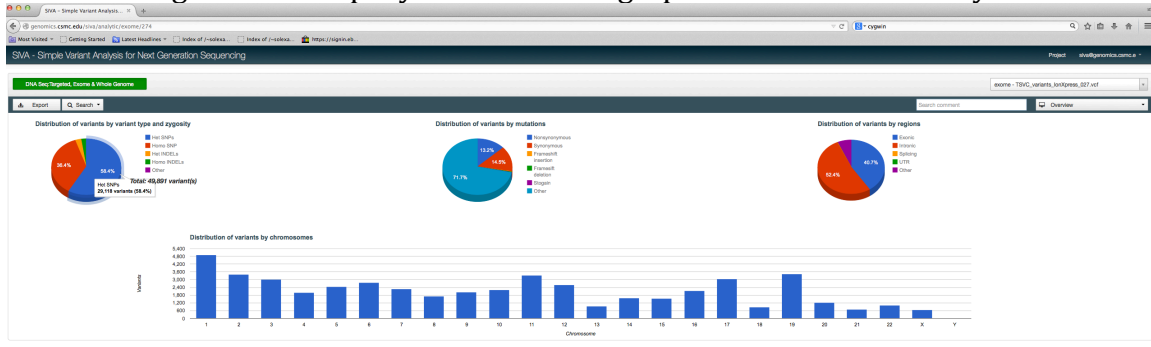
## Search Variants

Click on one sample name to view results and to search variants.

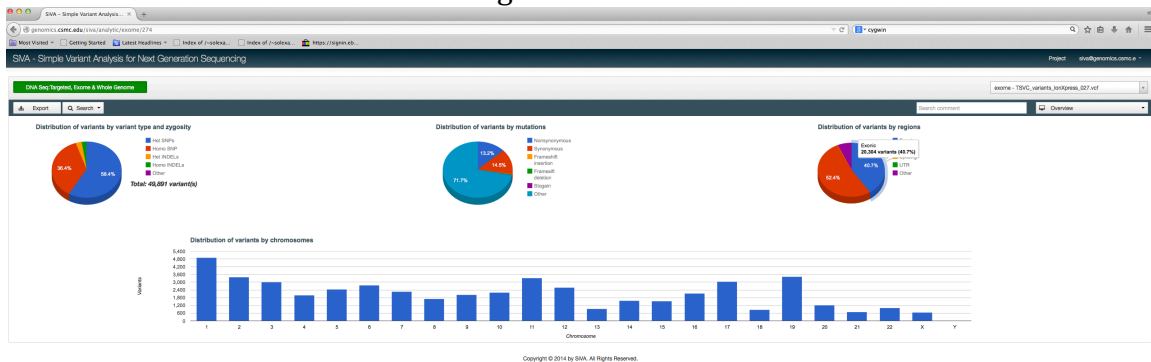


File Name	Created Date	Status	Size	Study Type	Source	Base	Action
TSVC_variants_bslpress_007.vcf	2014-08-14 10:27:01	Analyzed	22.67 MB	Cancer	Tissue	Other	<a href="#">View</a> <a href="#">Download</a>
TSVC_variants_bslpress_008.vcf	2014-08-14 10:27:41	Analyzed	23.12 MB	Cancer	Tissue	Other	<a href="#">View</a> <a href="#">Download</a>
TSVC_variants_bslpress_009.vcf	2014-08-14 10:27:58	Analyzed	28.22 MB	Cancer	Blood	Other	<a href="#">View</a> <a href="#">Download</a>

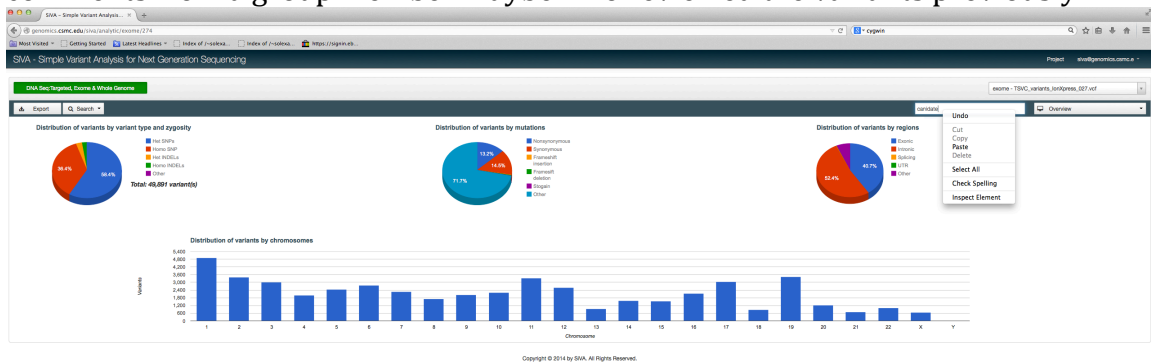
After clicking on one sample you can view in graphical detail a summary of variants



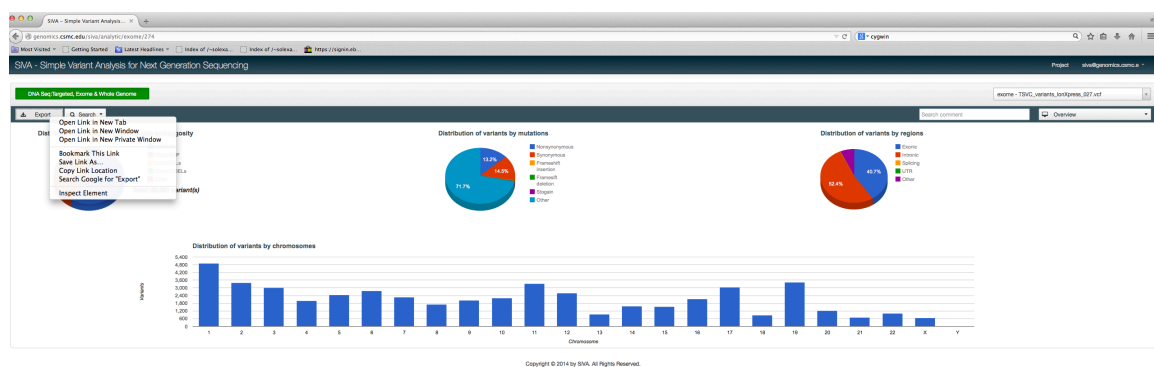
You can mouse over each section to get more information of the variants.



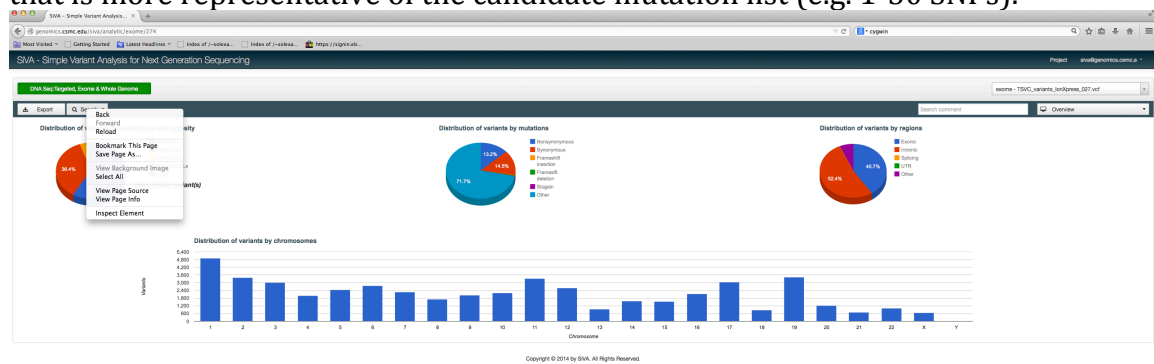
You can search for a Comment in the **Search** window above the Distribution of variants pie graph this is helpful if you are a principle investigator reviewing comments from a group member maybe who reviewed the variants previously.



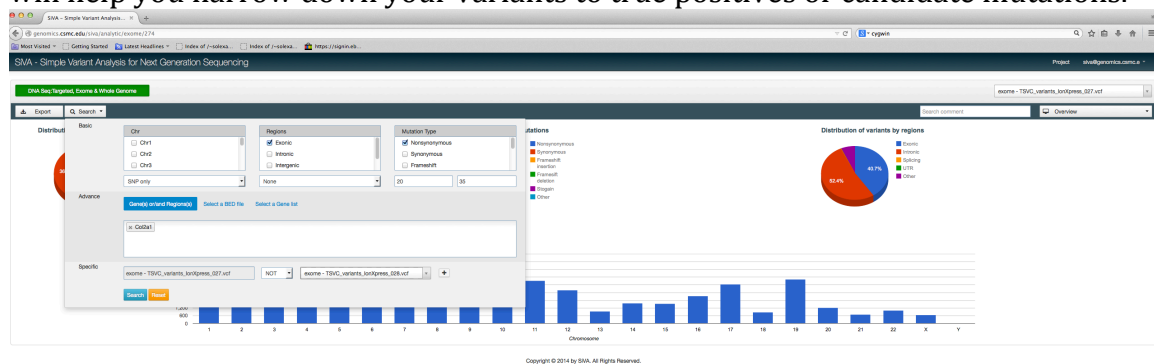
Once you upload the .VCF file and its annotated you can click **“Export”** on the far left (next to the **Search** key). Or you can search it and export only the searched values.



You can then search or filter all the variants in the sample (e.g. 50K) to something that is more representative of the candidate mutation list (e.g. 1-50 SNPs).

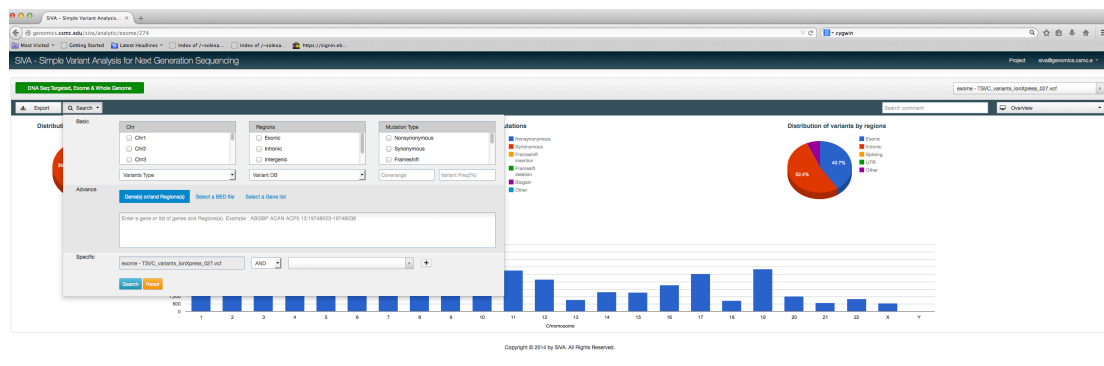


After you click search, a Search tab opens (see below) that includes many fields that will help you narrow down your variants to true positives or candidate mutations.



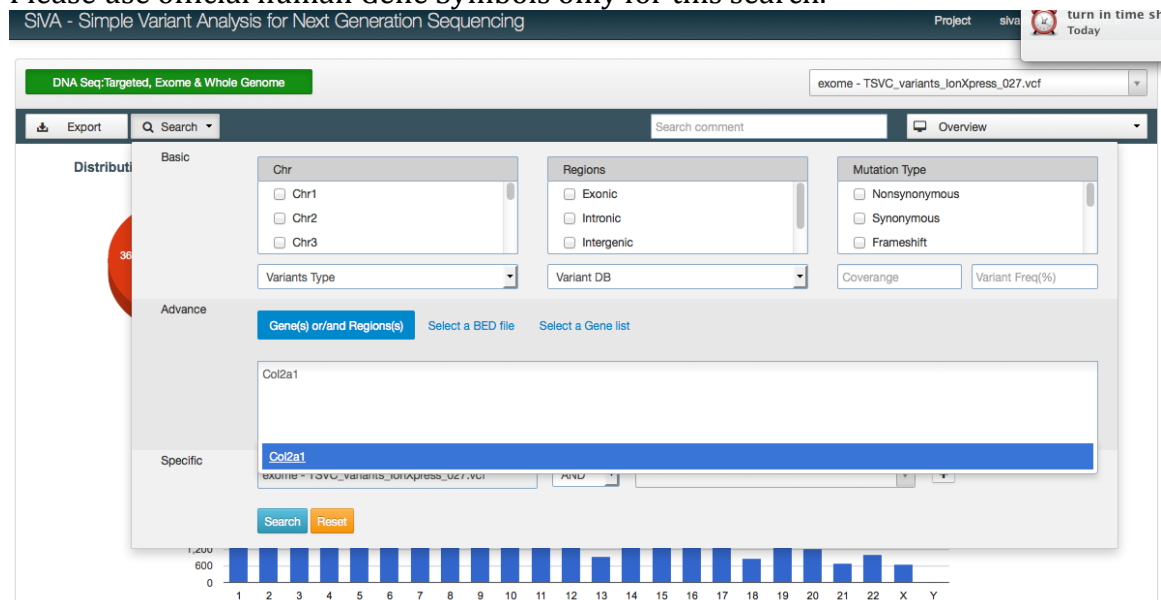
## Basic Search

In the search Table, select or enter reasonable filters that will result in a query that selects for only candidate mutations. For Example, in the case below we are searching a sample for a mutation that is inherited and not found in a sibling. Specifically, in the search tab we have selected to narrow our search to Exons (by checking “Exons” in “Regions” box), only Non-Synonymous variants (by checking “non-synonymous” in “Mutation type” box). Then narrowed the candidates by only looking at rare SNPs, by pulling down the “variants type” and “variant db” so that we selected “SNPs” (not indels) and “none” (for not in dbSNP). Then we enriched or selected for the heterozygous variants that represented at least 35 percent of the reads with at least a total coverage of 20.



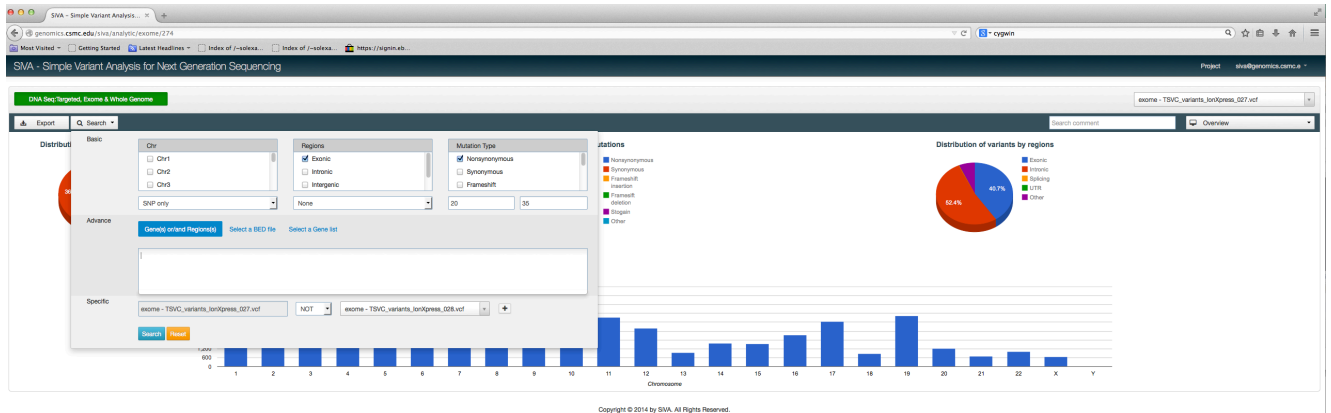
## Advanced SNP filtering/Analysis

Under the “Basic” SNP Filtering section is the “Advanced” SNP Search Window. In this window you can construct a list of genes that are routinely searched and upload the list to your account for future searches. This is helpful for routine searches. If you have 1 or 2 genes you just want to check then you can type the names of the genes in the window (As seen below in the screenshot where “Col2a1” was entered). Please use official human Gene Symbols only for this search.



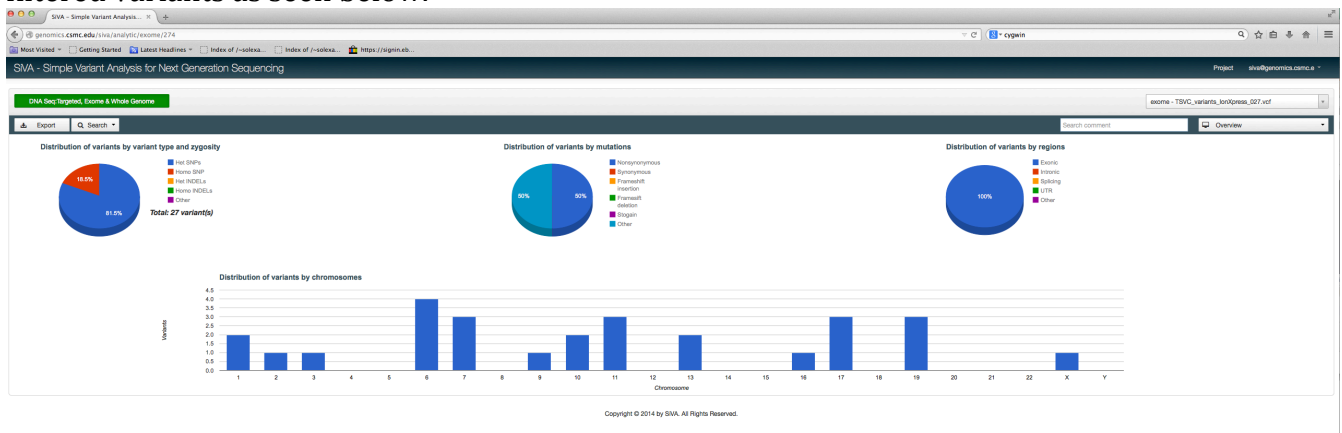
## Variants “Specific” to one sample

Under the Advanced SNP filtering there is a “**Specific**” filtering window to select for Variants that are uniquely found in one sample and/or not another sample. The Boolean query can be added sequentially with as many samples that are present in the users account. In this case, we selected not found in another sample to find a mutation that was unique and rare and not found in another sequencing sample. This can be very useful for separating out rare mutations from batch sequencing errors that may be in unrelated samples, as much as it is used to identify “**denovo**” mutations found in an unaffected parent for example.



## Dynamic Graphical user interfaces rapidly shows Summary of Filtering

After filtering the graphical interface is updated with the information from the filtered variants as seen below.



## Variant Annotation Details to further identify true candidate mutations.

After viewing the graphic summary, one can research the variant annotation in more detail by selecting the tab on the far right labeled **“Overview”**. When the tab is pulled down there reveals multiple annotation tabs.

### A. Variant Detail tab.

If you select Variant Detail you will see the following information

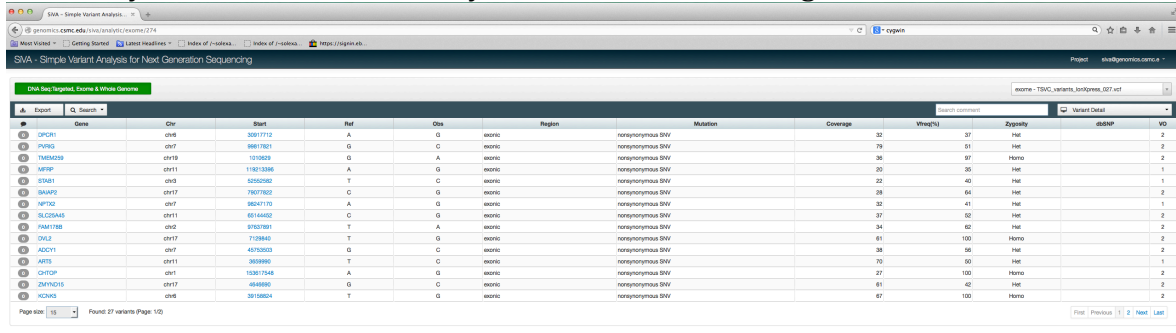


Table with 14 columns: Gene, Chr, Start, Ref, Obs, Region, Mutation, Coverage, VQ, Zygosity, and VO. The table lists 17 variants across various genes including DPCH1, PVRG, TMEM259, MPRP, ENAH, SNAI2, IL23RA5, TMEM178B, DLX3, ADCY1, ARS, COTOP, ZNF425, and KCMAS.

Gene	Chr	Start	Ref	Obs	Region	Mutation	Coverage	VQ	Zygosity	VO
DPCH1	chr6	20017712	A	G	exonic	heterozygous SNV	32	37	Het	2
PVRG	chr7	89471821	G	G	exonic	heterozygous SNV	79	91	Het	2
TMEM259	chr19	1101629	G	A	exonic	heterozygous SNV	36	97	Homo	2
MPRP	chr11	11813386	A	G	exonic	heterozygous SNV	20	35	Het	1
ENAH	chr3	5202082	T	G	exonic	heterozygous SNV	22	40	Het	1
SNAI2	chr17	7957352	C	G	exonic	heterozygous SNV	36	64	Het	3
MPR2	chr7	85247170	A	G	exonic	heterozygous SNV	32	41	Het	1
IL23RA5	chr11	8214482	C	G	exonic	heterozygous SNV	37	52	Het	2
TMEM178B	chr2	8762191	T	A	exonic	heterozygous SNV	34	62	Het	2
DLX3	chr17	7129842	T	G	exonic	heterozygous SNV	81	100	Homo	2
ADCY1	chr7	4570553	G	G	exonic	heterozygous SNV	38	56	Het	2
ARS	chr11	3639190	T	G	exonic	heterozygous SNV	70	50	Het	1
COTOP	chr1	10381768	A	G	exonic	heterozygous SNV	27	100	Homo	2
ZNF425	chr17	4648890	G	G	exonic	heterozygous SNV	61	42	Het	2
KCMAS	chr6	20158524	T	G	exonic	heterozygous SNV	67	100	Homo	2

B. The second tab is **“Annotation and Databases”** and provides rich annotation of the variant frequency and AA changes.

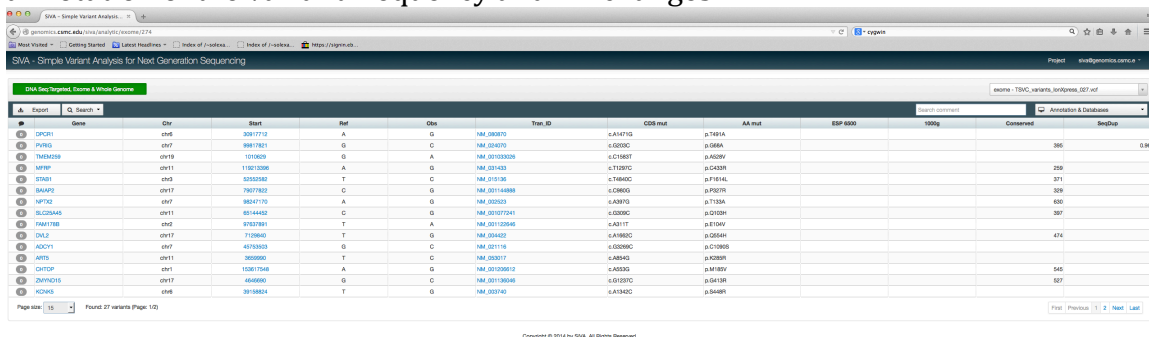


Table with 14 columns: Gene, Chr, Start, Ref, Obs, Trn, ID, CDS mut, AA mut, ESP 6500, 1000G, Conserved, and RefSeq. The table lists 17 variants across various genes including DPCH1, PVRG, TMEM259, MPRP, ENAH, SNAI2, IL23RA5, TMEM178B, DLX3, ADCY1, ARS, COTOP, ZNF425, and KCMAS.

Gene	Chr	Start	Ref	Obs	Trn	ID	CDS mut	AA mut	ESP 6500	1000G	Conserved	RefSeq
DPCH1	chr6	20017712	A	G		NA_020715	c.A1471G	p.T59A				
PVRG	chr7	89471821	G	G		NA_016170	c.C200C	p.S68A				360
TMEM259	chr19	1101629	G	A		NA_0103326	c.C158T	p.A52V				259
MPRP	chr11	11813386	A	G		NA_0112015	c.T120T	p.C43V				271
ENAH	chr3	5202082	T	G		NA_011336	c.T64G	p.T21A				329
SNAI2	chr17	7957352	C	G		NA_0214488	c.C86G	p.P27R				630
MPR2	chr7	85247170	A	G		NA_022523	c.A29T	p.T135A				397
IL23RA5	chr11	8214482	C	G		NA_0202791	c.C200C	p.C200R				476
TMEM178B	chr2	8762191	T	A		NA_01122546	c.A31T	p.E104V				
DLX3	chr17	7129842	T	G		NA_024421	c.A188C	p.D54H				
ADCY1	chr7	4570553	G	G		NA_011716	c.C200G	p.C196G				
ARS	chr11	3639190	T	G		NA_020217	c.A484S	p.K289R				
COTOP	chr1	10381768	A	G		NA_01038512	c.A53G	p.M180V				545
ZNF425	chr17	4648890	G	G		NA_0212648	c.A123T	p.S41R				827
KCMAS	chr6	20158524	T	G		NA_020170	c.A134G	p.S44H				

C. The Next tab **Functional Prediction** includes the results from several prediction algorithms on the effect of the variant on the mutation.

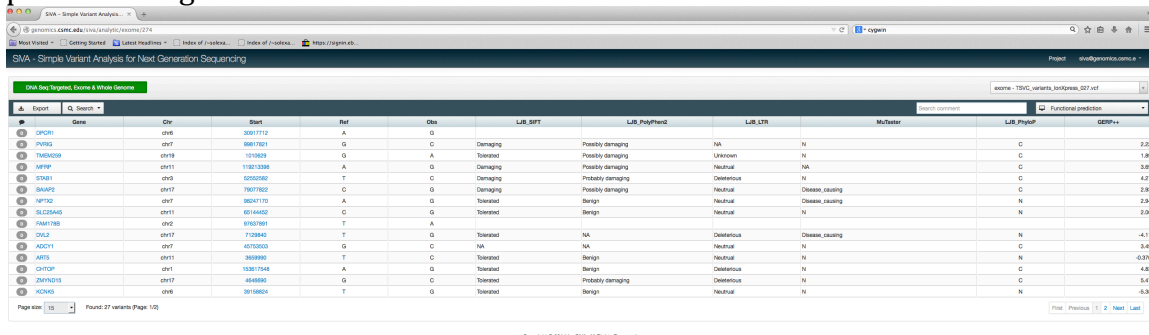


Table with 14 columns: Gene, Chr, Start, Ref, Obs, LRS\_SIFT, LRS\_PolyPhen2, LRS\_LRT, Mutator, LRS\_PhysP, and GERP. The table lists 17 variants across various genes including DPCH1, PVRG, TMEM259, MPRP, ENAH, SNAI2, IL23RA5, TMEM178B, DLX3, ADCY1, ARS, COTOP, ZNF425, and KCMAS.

Gene	Chr	Start	Ref	Obs	LRS_SIFT	LRS_PolyPhen2	LRS_LRT	Mutator	LRS_PhysP	GERP
DPCH1	chr6	20017712	A	G	Damaging	Probably damaging	NA	N	C	2.22
PVRG	chr7	89471821	G	G	Tolerated	Probably damaging	Unknown	N	C	1.99
TMEM259	chr19	1101629	G	A	Tolerated	Probably damaging	Unknown	N	C	2.89
MPRP	chr11	11813386	A	G	Damaging	Probably damaging	Neutral	NA	C	4.27
ENAH	chr3	5202082	T	G	Damaging	Probably damaging	Damaging	N	C	2.88
SNAI2	chr17	7957352	C	G	Damaging	Probably damaging	Neutral	Disease_causi	C	2.88
MPR2	chr7	85247170	A	G	Tolerated	Benign	Neutral	Disease_causi	N	2.84
IL23RA5	chr11	8214482	C	G	Tolerated	Benign	Neutral	N	N	2.05
TMEM178B	chr2	8762191	T	A	Tolerated	NA	Neutral	Disease_causi	N	-4.11
DLX3	chr17	7129842	T	G	Tolerated	NA	Neutral	N	C	3.49
ADCY1	chr7	4570553	G	G	NA	Neutral	Neutral	N	C	-0.76
ARS	chr11	3639190	T	G	Tolerated	Benign	Damaging	N	C	4.42
COTOP	chr1	10381768	A	G	Tolerated	Benign	Damaging	N	C	5.47
ZNF425	chr17	4648890	G	G	Tolerated	Probably damaging	Damaging	N	C	-0.35
KCMAS	chr6	20158524	T	G	Tolerated	Benign	Neutral	N	N	

After viewing the variants list, one can right click on the hypertext gene symbol to learn more information about the gene available in different databases like OMIM, ClinVar, etc. As seen below in the pop up menu in the screen shot.

Finally, one can drill down into more variant detail by looking at the position of the variant in respect to other genomic data in UCSC genome browser. By left clicking on the hypertext position of the variant (as seen below in the screen shot) one can open a new window that opens at the variant position in the human genome in the UCSC genome browser.

The screenshot shows the UCSC Genome Browser interface. The top navigation bar includes links for 'Genes', 'RefSeq', 'Ensembl', 'NCBI', 'UCSC', 'GenBank', 'EMBL', 'DDBJ', 'PDB', 'RCSB', 'EMBL', 'NCBI', 'UCSC', 'GenBank', 'EMBL', 'DDBJ', 'PDB', 'RCSB'. The main content area displays a genomic track for chromosome 10, with a variant at position 110310. The variant is a G to T substitution. The browser interface includes a search bar, a list of tracks (Genes, RefSeq, etc.), and a detailed view of the variant. A tooltip is visible over the variant position, showing a link to 'Open Link in New Window'.

[illegible]